METHOD AND COMPUTER PROGRAM PRODUCT FOR DRUG DISCOVERY USING WEIGHTED GRAND CANONICAL METROPOLIS MONTE CARLO SAMPLING

ABSTRACT OF THE DISCLOSURE

A method and computer program product for modeling a system that includes a protein and a plurality of fragments in order to identify drug leads is presented. The basis of the method is a weighted Metropolis Monte Carlo approach for sampling the Grand Canonical ensemble. This method distinguishes itself from an energy minimization approach in that it provides fragment distributions which are consistent with thermal fluctuations at physiologically relevant temperatures. The weighted Metropolis Monte Carlo scheme performs a quasi-uniform sampling of all regions of interest on the protein, and, in this way, enables to resolve the wide range in densities of the thermodynamic distribution which could not be achieved by a non-weighted Metropolis scheme. Making use of the properties of the Grand Canonical ensemble, the affinity of fragments for different regions on the protein surface can be efficiently computed. A protein binding site is then identified as a region with high affinity for multiple fragments with a diverse set of physico-chemical properties. Within a binding site, assembly of fragments into drug leads is finally carried out based on binding affinity of the different fragments, on geometric proximity, and a variety of rules by which organic fragments may bond together.